DOCUMENT RESUME

ED 047 309 AL 002 751

AUTHOR	Guyette, Thomas W.; And Others
TITLE	Reconstruction: A Necessary Stage in the Analysis of Verbal Interaction.
INSTITUTION	Missouri Univ., Columbia. Center for Research in Social Behavior.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Feb 71
NOTE	16p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, N.Y., February 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Age Differences, Caucasian Students, *Classroom Communication, *Content Analysis, Discourse Analysis, Grammar, Information Theory, *Interaction Process Analysis, Negro Students, *Racial Differences, Research Methodology, Semantics, *Sociolinguistics, Student Evaluation
IDENTIFIERS	Paraphrase

ABSTRACT

This paper argues that traditional techniques of content analysis which have been applied to classroom verbal interaction studies fail to account for certain types of implicit information, and presents a technique, called "reconstruction," that provides for the inclusion of such information in a content analysis. It is argued that if identical messages (paraphrases) can contain varying amounts of overt information then the case for implicit information and thus reconstruction is supported. In other words, if individuals who interpret messages must supply information which is not overtly available, then this implicit information must be included to obtain frequency counts which accurately reflect the semantic preoccupations of the verbal interaction. The paper first discusses the existence of implicit information in a communicative event and a technique of reconstruction for discovering that information. The second part presents data to show that reconstruction makes a difference, and in particular that black and white classrooms differ in the degree to which they exhibit implicit information in their discourse. These differences are felt to be substantive. It is also felt that, without such reconstruction, content analyses of classroom discourse generate distorted views of the concerns of participants. See also AL 002 750 and AL 002 752-753. (Author/FWB)

ED047309

RECONSTRUCTION: A NECESSARY STAGE IN THE ANALYSIS OF VERBAL INTERACTION *

by

Thomas W. Guyette
Marvin D. Loflin
Nancy Barron
Marjorie Marlin

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

AL 002 751

* The research leading to the results reported here have been supported by funds from the National Science Foundation, Grant No. GS3232 and by Institutional Research Funds to the Center for Research in Social Behavior at the University of Missouri, Columbia, Missouri.

Certain analytic techniques devised for the analysis of verbal behavior claim to reflect the density of semantic information in a communicative event. These techniques usually involve some form of content analysis and are applicable to classroom verbal interaction studies. Unfortunately, most studies of classroom interaction do not meet the standards of this formal technique. In content analysis, words uttered by participants are assigned by the analyst to categories that represent their meanings. In linguistic terms, content analysis involves placing lexical items into semantic categories. The frequency of words in these categories is then assumed to reflect the semantic preoccupations of the classroom lesson under analysis.

This paper argues that traditional techniques of content analysis have failed to take into consideration certain types of implicit information in the analysis of verbal behavior. It is the purpose of the paper to present and justify a technique, called reconstruction, that provides for the inclusion of implicit information in a content analysis. The argument put forth claims that if identical messages (that is, paraphrases) can contain varying amounts of overt information then the case for implicit information and thus reconstruction is supported. In other words, if individuals who interpret messages must supply information which is not overtly available, then this implicit information must be included to obtain frequency counts which accurately reflect the semantic preoccupations of the verbal interaction. The first part of the paper duscusses the existence of implicit information in a communicative event and a

technique of reconstruction for discovering that information. In the second part of the paper, data are presented to show that reconstruction makes a difference, and in particular that black and white classrooms differ in the degree to which they exhibit implicit information in their discourse. These differences are substantive. Thus it will be shown that content analyses of classroom discourse generate distorted views of the concerns of participants unless reconstruction of the discourse is carried out.

Implicit Information and Reconstruction

Traditional content analysis has operated under the assumption that all information communicated through language is overt in the behavioristic sense of the word. That is, all meaning which is communicated is present in the form of overt lexical items which, when counted, reflect the density of various meanings in a text. Although this assumption is inadequate (as will be shown below) it does seem to be the case that much information in a communicative event is inferable from overt verbal structures. This is important because it allows systematic reconstruction of implicit information by personnel trained to cue off of these overt verbal behaviors. In other words, overt verbal behavior cues language users to the presence and meaning of information that is not explicit in the communicative event being examined.

Information present in a communicative event may be categorized loosely into three types. First, it can be explicit in the behaviorist sense (that is, semantic information can be overtly present in the communication). Second, it can be grammatically implicit (that is, redundant units are deleted according to grammatical rules in a language). Third, it can be contextually implicit (that is, not in the overt message but inferable from the context). The process of adding grammatically and contextually implicit information to an explicit text is referred to as reconstruction.

What kinds of judgments are involved in reconstruction? In grammatical terms, a text may exhibit implicit information at either intersentential (discourse) or intrasentential (sentence) levels. At the discourse level, implicit units are sentences. If implicit sentences at the discourse level are not inserted into the protocol, an inaccurate picture of the semantic content of the speech being studied will emerge. For example, the sentence

(1) George is no player.

is implicit in

(2) George has missed fourteen freethrows and Bill is no player, either.

Thus, we would reconstruct sentence (2) as sentence (3), where reconstructions of implicit information are given in parentheses.

(3) George has missed fourteen freethrows (and George is no player) and Bill is no player, either.

A discussion of the content items in (2) without consideration for (1) would lead to the conclusion that as much is being asserted about Bill as about George. But, a discussion of the content items of (3) would reflect more content concern for George than for Bill. In order for a content analysis to adequately reflect the actual relative proportions of semantically same lexical units in a message it must present frequency counts that include the implicit lexical elements in sentence (3).

Even more important to content analyses are the implicit lexical elements within sentences. Sentences which contain the same semantic content and which may be related to each other by one or more explicit transformational processes may be said to be paraphrases, that is, the same message. As stated earlier, it is important that the same message be represented by the same frequency of lexical units or the assumptions of content analysis are violated.

For example, in the sentence set (4),(5), and (6):

(4) The girl laughed.

(5) The girl is telling a joke.

(6) The girl laughed while telling a joke.

sentence (6) is a possible paraphrase of sentence (4) and (5). A frequency count in the text which contained sentences (4) and (5) would yield two occurrences of the 'feminine' noun, 'girl.' However, if the speaker adjoined the two sentences and deleted the redundant subject as in (6) a count would yield only one occurrence of the 'feminine' noun 'girl.' In the reconstructed sentence

(7) The girl laughed while (the girl) telling a joke.

both occurrences are retained. Therefore, the reconstructed sentence (7) must be judged as a more valid representation of the content than sentence (6) if textual content is determined from frequency counts of word classes.

Sentences derived through conjunctive reduction (8) also provide examples of intrasentential implicitness. For example, sentence (8) is reconstructed as (9).

(8) The boys and the girls romped all day.

(9) The boys (romped all day) and the girls romped all day.

Once again, we find in a reconstructed protocol increased frequency of content elements.

Contextually implicit information must also be added in order to obtain a more accurate representation of the meaning in the text. Contextual implicits are structures which are inferred from the context of the communicative event.

In the sequence

Teacher: (10) Do you know the answer?

Pupil: (11) Yeah.

sentence (11) contains information which is not recorded by content analysis using unreconstructed text. The reconstructed sentence

(12) Yeah (I know the answer).

does contain the information that is inferred by the context of the interaction. Note that the reconstructed sentence (12) differs little from the first sentence (10). This is because contextually implicit structures are 'carried down' from preceding overt utterances. The response 'yeah' is bound to the context. Stated without the context it does not have the same meaning it does in a sequence of this sort.

A second type of contextually related implicit information concerns pronouns and their referents. Referents are real-world tokens. The problem of identifying successive repetitions of speech forms as referentially representing the same real-world tokens is considered a matter of providing implicit contextual information. An inspection of a sample of the speech used in the classroom may show that students names are not used. This fact notwithstanding, the pronoun <u>you</u> may appear 53 times in the sample. In order to determine which you-forms are to be associated with which students it will be necessary to go to the actual social scene and pair you-forms with the real-world tokens being referenced. Only by referencing pronouns is it possible to achieve an accurate representation of the semantic information which is communicated in the discourse.

A third kind of contextually implicit information in verbal interaction is to be found in speech forms which themselves have referents which also are speech forms in the communicative event. As the interaction proceeds, a speaker may make reference to ideas that are not in the immediate time frame of the interaction. Language is equipped with metalinguistic terms which make

it possible to displace and reallocate chunks of discourse which may vary from very short sub-word morphemic structures such as the comparative 'er' or very long stretches such as 'all that was said at the United Nations on the 24th of January 1970.' In addition to grammatical terms such as verb and sentence, other terms which reference verbal behavior, written or spoken, must be considered. For example, a teacher may ask, "What do we learn about the customs of other people from this story?" The word story references a body of latent verbal behavior which when read becomes the potential focus for discussion. Teachers obviously use forms of this sort to direct attention to communicative events that are displaced in space and time.

Analyses of verbal behavior often employ statistics. Statistical comparison is grounded on certain measurement assumptions. A basic assumption is that one occurrence of the unit we are counting is weighted equally with any other occurrence of the unit. An examination of sentence (1) through (12) illustrates that unreconstructed sentences are biased by underrepresentation due to the exclusion of implicit information. Reconstructed sentences (3,7, 9 and 12), however, yield an unbiased representation of the words which classify into semantic categories. From this it is only reasonable to conclude that content analyses which use unreconstructed texts do not meet basic measurement assumptions.

If the preceding discussion argues effectively for the presence of implicit information in a text then it must be apparent that frequency counts based on texts which exclude a significant portion of the countable elements are inadequate. To correct this inadequacy we propose that implicit information be included in the text preparation stage of the semantic analysis (see Loflin and Barron, 1970). Including reconstructed material increases the validity of

the assumption that the meaning in a text is reflected by frequency counts of the lexical units. Statistical descriptions of content arrived at with reconstruction are based on a foundation of increased validity; statistical descriptions arrived at without it are less valid.

Finally, something must be said about the frequency of implicit lexical elements in various contexts. If implicit information occurred in approximately the same distribution as explicit information, then in large samples there would be little reason to reconstruct as meaning discrepancies due to implicits would become negligible. However, if the opposite hypothesis is true(that is, if the distribution of implicit information is not the same as the distribution of explicit information) then reconstruction becomes necessary to correct the semantic bias of unreconstructed texts.

This hypothesis will be examined as it related to the implicit and explicit distribution of case forms. Case is a linguistic variable that yields information about the semantic relationship of a noun phrase to its main verb. In other words, the case of a noun phrase is determined by an examination of the semantic relationship which exists between a noun phrase and the verb in a simplex sentence. Consider sentence (13).

(13) John hit Larry.

The subject 'John' is the agent of the action of the verb 'hit.' The object 'Larry' however exhibits a different semantic relationship in that it is the receiver of the action of the verb. The subject of (13) 'John' is an example of an agentive case noun phrase. The relationship between noun phrase and verb is of the type 'agent-action.' The object of (13) 'Larry' is an example of a dative case noun phrase. The relationship between the noun phrase and the verb is of the type 'action-receiver.' In the study to be reported below, fifteen

cases were used to classify the semantic relationships between noun phrases and verbs in the examples of the classroom discourse analyzed.

An analysis of the distribution of case in a communicative event is a form of content analysis. In case analysis semantic information is counted and placed into categories which are assumed to represent a level of semantic concern in the text. For example, if the classroom speech under analysis contained a large amount of agentive case noun phrases, then it would be logical to infer that one of the central themes of the discussion concerned animate beings who were doing or acting. In the same way, if the speech contained a preponderance of dative cases, it would indicate a semantic concern with animate beings who were affected by the action of the verb.

Results

Two sets of data are to be presented. The first concerns the uses of case in the speech of eight teachers in classroom interaction situations. One interesting observation derived from these data concerns the amount of implicit information in a communicative event. Of all cases used approximately 45% were implicit. This suggests a much greater amount of implicit information than is commonly expected. For all fifteen cases a chi-square analysis was used to determine if significant differences existed between the distributions of implicit and explicit use of case. The results confirmed the hypothesis ($p<.001$) that significant differences do exist between the two distributions.

Which cases are more likely to receive implicit use? Data answering this question appear in Table 1 in the form of percentages of explicit and implicit use As will be seen in this Table, it was found that there was a greater percent of agentive case noun phrases in implicit use than in explicit use ($p<.01$). These results suggest that information concerning animate beings

who are agents of an action is under-represented in a content analysis which counts only surface or explicit meanings. It was also found that participative case noun phrases (that is, animate beings which participate in the action of an internal state psychological verb) were more often used in explicit speech than in implicit speech ($p<.02$). Thus, the participative case noun phrases are over-represented in unreconstructed text. These differences were large enough to call for a reinterpretation of the data given the addition of implicit information. In this example,the participative case was the second and agentive case the third most frequently used case in the explicit case distribution. However, with the addition of implicit information the agentive case becomes the second and the participative the third most frequently occurring case category. That is, the addition of implicit information forces a new interpretation of the results.

In addition, the dative case has more implicit realizations than explicit ($p<.10$) and the location case has more explicit realizations than implicit ($p<.10$). This indicates that information concerning animate beings who are being acted upon is under-represented in the explicit distribution. The converse is true of the location case (that is, it is over represented in the explicit distribution.).

A final difference in the distribution of implicit and explicit information occurs in the manner case. These noun phrases relate to the main verb in that they describe or tell in what manner the action of the verb occurred (for example, He walked <u>with a limp</u>.). This case is under-represented in the implicit distribution. It occurs implicitly approximately one time in four explicit occurrences. The distribution of the manner case is significantly different ($p<.01$) between implicit and explicit distributions.

In general, then, it appears that for these data there are significant differences between the implicit and explicit distributions of the meaning variable case. The implications of this are that texts which are not reconstricted will contain meaning biases large enough in some instances to call for a reinterpretation of results.

It was also hypothesized that there would be differences in the amounts of information used by different social groups. For example, such independent variables as race of classroom should significantly affect the amount of implicit information which occurred. This hypothesis was tested on a second set of data derived from fifteen ten-minute sections of audiovideo tapes taken from black and white classrooms. In contrast with the results reported above, these latter data included the responses of both pupils and teachers.

Significant differences were found between amounts of implicit information in the speech of inner city blacks and suburban whites (see Table 11 and Figure 1). As regards major lexical categories (that is, words which are not proforms), black classrooms exhibited a higher rate of implicit information ($p < .01$) than white classrooms. This is to say, persons in black classrooms delete more information from their speech than do persons in white classrooms. However, this difference was not found at all grade levels. Black classrooms features more implicit information at the first and sixth grade levels but not eleventh grade level. In short, the differential use of implicit information appeared to disappear in the eleventh grade. These results suggest that content analyses which use unreconstructed data taken from the speech of interracial groups are biased toward the speech of white, middle class participants. This follows logically from the empirical results which show that blacks delete more information from their speech than do whites.

Differences were also found between black and white classrooms in the relative frequency of implicit proforms (see Table III and Figure 11). Proforms include pronouns like 'he' or 'she' and general category words such as 'thing.' Significant differences were found between black and white first (p .01), sixth ($p<.01$), and eleventh ($p<.01$) grade classrooms. In all cases black classrooms exhibited higher percentages of implicit proforms than white classrooms, and these differences did not disappear as pupils became older. These results also suggest that content analysis which use unreconstructed data taken from the speech of interracial groups are biased toward the speech of white, middle class participants.

Quite apart from black-white differences, it was found that eleventh ($p<.01$), sixth ($p<.01$) and first ($p<.01$) grade classrooms deleted significantly larger amounts of major lexical categories than proforms. This indicates that information transmitted in the form of a proform is less likely to be deleted than corresponding information that is transmitted through major lexical categories.This has implications for content analyses which referent only explicit proforms. In such analyses, proform information is over-represented in the explicit distribution.

It was also found that use of proforms varies from grade level to grade level. Proform use is higher in the first grade than in the eleventh grade ($p<.01$) with sixth grade exhibiting intermediate usage. These results indicate that a significantly greater amount of information is transmitted through proforms in lower grades. This has implications for classroom interaction research which is involved in inter-grade-level-comparisons.

The fact that black classroom exhibit a significantly larger percentage of implicit speech than white classrooms suggests that black speech is more

context bound than white speech. That is, in order to interpret any given sentence in a black classroom the probability is greater that access to the context will be necessary. The converse is true in white classrooms where it is more likely that any given sentence could be interpreted without reference to the context. If we assume that an educational goal is to make utterances less context restricted, and more explicit, then we are forced to the conclusion that because black students do not adhere to this norm they will probably be negatively evaluated by their teachers.

Summary

In summary, it has been shown that individuals who interpret messages must supply information which is not available in overt verbal behavior. It was argued that this information must be included in analyses which seek to discover the semantic preoccupations of a verbal interaction. This argument was supported by empirical data which indicate that the distribution of implicit case noun phrases differs from that of explicit or surface case noun phrases. In addition, substantive differences were found to exist between the amount of implicit information in black and white classroom speech. Garde level differences were also found. These suggest that lower grade level classrooms use more proforms than upper grade level classrooms. Finally, it was found that proforms are less likely to be deleted than major lexical categories. In general, the results suggest that the distribution of implicit information is affected by social psychological variables such as race and grade level. It must be concluded,then, that the distribution and quantity of implicit and explicit information which occur in natural speech can no longer be assumed to be the same. It is because this assumption can no longer be made that reconstruction is a necessary stage in the analysis of verbal interaction.

TABLE I
DIFFERENCES IN THE PERCENTAGE OF IMPLICIT AND EXPLICIT CASES

CASE	PERCENT EXPLICIT	IMPLICIT	DIFFERENCE	SIGNIFICANCE
Agentive	17.7	23.6	5.9	$p < .01$
Participative	18.1	13.7	4.4	$p < .02$
Dative	8.8	11.0	2.2	$p < .10$
Location	8.4	6.0	2.4	$p < .10$
Manner	2.1	.4	1.7	$p < .01$

TABLE II
DIFFERENCES IN THE PERCENT OF IMPLICIT MAJOR LEXICAL CATEGORIES IN THE SPEECH OF BLACK AND WHITE CLASSROOMS BY GRADE LEVEL

GRADE LEVEL	PERCENT IMPLICIT BLACK	WHITE	DIFFERENCE	SIGNIFICANCE
First	52.5	42.6	9.9	$p < .01$
Sixth	52.7	38.4	14.3	$p < .01$
Eleventh	44.9	46.4	1.5	$p = ns$

TABLE III
DIFFERENCES IN THE PERCENT OF IMPLICIT PROFORMS IN THE SPEECH OF BLACK AND WHITE CLASSROOMS BY GRADE LEVEL

GRADE LEVEL	PERCENT IMPLICIT BLACK	WHITE	DIFFERENCE	SIGNIFICANCE
First	46.2	36.5	9.7	$p < .01$
Sixth	44.9	35.3	9.6	$p < .01$
Eleventh	45.6	35.9	9.7	$p < .01$

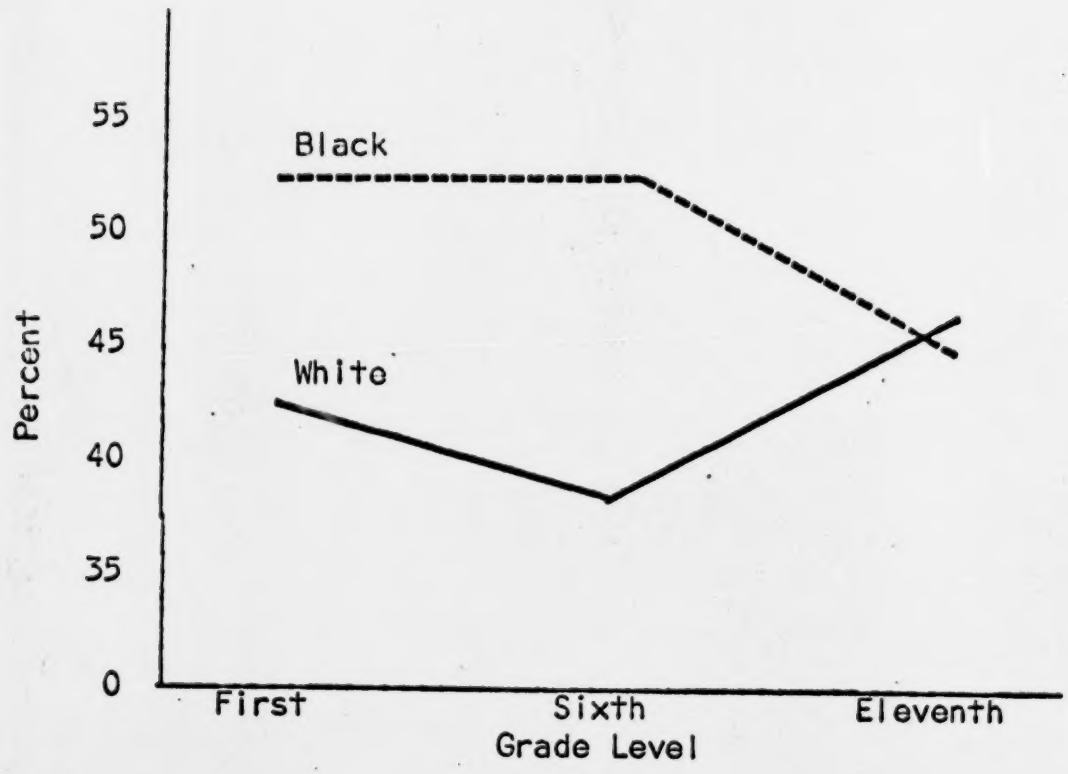

Figure I
Percentage of Implicit Major Lexical Categories in the Speech of Black and White Classrooms by Grade Level

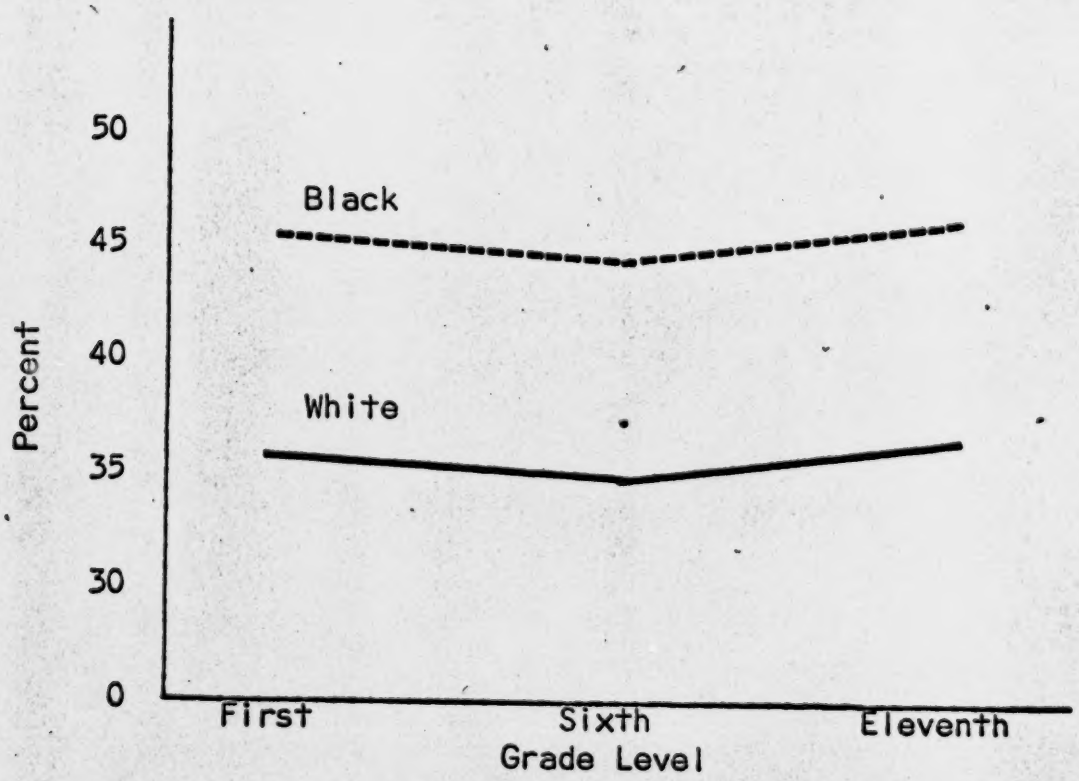

Figure II
Percentage of Implicit Proforms in the Speech of Black and White Classrooms by Grade Level

REFERENCES

Loflin, Marvin D., and Nancy Barron. "Reconstructing Classroom Behavior: A Manual." Technical Report No. C132. The Classroom Interaction Project Series. Center for Research in Social Behavior, University of Missouri--Columbia, 1970.